

대한민국 특허청
KOREAN INTELLECTUAL
PROPERTY OFFICE

별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto
is a true copy from the records of the Korean Intellectual
Property Office.

출원번호 : 10-2002-0075407
Application Number

출원년월일 : 2002년 11월 29일
Date of Application NOV 29, 2002

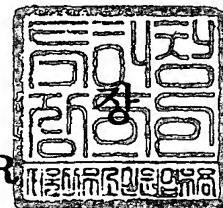
출원인 : 삼성전자주식회사
Applicant(s) SAMSUNG ELECTRONICS CO., LTD.



2003 년 04 월 25 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0018
【제출일자】	2002.11.29
【국제특허분류】	H03M
【발명의 명칭】	개선된 심리 음향 모델을 이용한 디지털 오디오 부호화 방법과 그 장치
【발명의 영문명칭】	Method for encoding digital audio using advanced psychoacoustic model and apparatus thereof
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	1999-009556-9
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2000-002816-9
【발명자】	
【성명의 국문표기】	마쓰 마누
【성명의 영문표기】	MATHEW,Manu
【주소】	경기도 수원시 팔달구 매탄3동 성일아파트 202동 205호
【국적】	IN
【우선권주장】	
【출원국명】	US
【출원종류】	특허
【출원번호】	60/000,000
【출원일자】	2002.10.30
【증명서류】	미첨부
【심사청구】	청구

【취지】

특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인

이영필 (인) 대리인

이해영 (인)

【수수료】

【기본출원료】

20 면 29,000 원

【가산출원료】

20 면 20,000 원

【우선권주장료】

1 건 26,000 원

【심사청구료】

34 항 1,197,000 원

【합계】

1,272,000 원

【첨부서류】

1. 요약서·명세서(도면)_1통 2. 우선권증명서류 및 동 번역문_1통

【요약서】**【요약】**

본 발명은 개선된 심리 음향 모델을 사용한 디지털 오디오 부호화 방법에 관한 것으로서, 본 발명에 따른 오디오 데이터 부호화 방법은 입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 단계와, 결정된 윈도우 타입에 따라 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 단계와, 결정된 윈도우 타입을 이용하여 입력 오디오 신호로부터 FFT 스펙트럼을 생성하는 단계와, 생성된 CMDCT 스펙트럼 및 FFT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 단계를 포함한다.

【대표도】

도 2

【명세서】**【발명의 명칭】**

개선된 심리 음향 모델을 이용한 디지털 오디오 부호화 방법과 그 장치 {Method for encoding digital audio using advanced psychoacoustic model and apparatus thereof}

【도면의 간단한 설명】

도 1은 종래의 MPEG 오디오 부호화 장치를 도시하는 블록도

도 2는 본 발명의 일 실시예에 따른 MPEG 오디오 부호화 장치를 도시하는 블록도

도 3은 본 발명에 따른 윈도우 스위칭 알고리즘에 사용되는 천이 신호 검출 방식을 도시하는 도면

도 4는 본 발명에 사용되는 윈도우 스위칭 알고리즘을 도시하는 플로우 차트

도 5는 본 발명에 따른 서브-밴드 스펙트럼으로부터 전체 스펙트럼을 구하는 방식을 도시하는 도면

도 6은 본 발명의 일 실시예에 따른 MPEG 오디오 부호화 방법을 도시하는 플로우 차트

도 7은 본 발명의 일 실시예에 따른 MPEG 오디오 부호화 장치를 도시하는 블록도

도 8은 본 발명의 일 실시예에 따른 MPEG 오디오 부호화 방법을 도시하는 플로우 차트

【발명의 상세한 설명】**【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<9> 본 발명은 디지털 오디오 부호화를 위한 부호화 방법 및 그 장치에 관한 것으로서, 특히 개선된 심리 음향 모델을 사용하여 음질의 저하 없이 부호화 방법 및 그 장치에 소요되는 계산량을 감소시키고, 복잡도를 감소시키기 위한 부호화 방법 및 장치에 관한 것이다.

<10> MPEG 오디오 부호화기는 부호화 중 생성된 양자화 잡음을 청취자가 지각(perception)하지 않도록 함과 동시에 높은 압축율(high compression rate)을 실현한다. MPEG에서 표준안을 정한 MPEG-1 오디오 부호기는 오디오 신호를 32kbps에서 448kbps의 비트율로 부호화한다. MPEG-1 오디오 규격은 부호화를 위한 3개의 다른 알고리즘을 갖는다.

<11> MPEG-1 부호화기는 레이어 1, 2, 3라고 하는 세가지의 모드를 가지고 있다. 레이어 1은 가장 기본적인 알고리즘을 구현하고, 레이어 2 및 3은 레이어 1이 개선된(enhanced) 것이다. 높은 레이어 일수록 고품질과 고 압축률이 실현되는 반면 하드웨어의 규모는 커진다.

<12> MPEG 오디오 부호화기는 신호의 지각적 중복성을 줄이기 위해 인간의 청각 특성을 잘 반영하는 심리 음향 모델을 사용한다. MPEG에서 표준안을 정한 MPEG 1과 MPEG 2는 인간의 지각 특성을 반영하고, 지각적 중복성을 제거하여 부호화 후에도 좋은 음질이 유

지될 수 있도록 심리 음향 모델 (psychoacoustic model)을 이용하는 지각 부호화 방식을 채택하고 있다.

<13> 지각 부호화(perceptual coding) 방식은 사람의 심리 음향 모델을 분석하여 적용한 기법으로서, 최소 가청 한계(threshold in quiet)와 마스킹 효과(masking effect)를 이용한다. 마스킹 효과는 큰 음에 의해 어떤 임계 값 이하의 작은 음이 가려지는 현상을 말하며, 이와 같이 같은 시간에 존재하는 신호간의 마스킹을 주파수 마스킹(frequency masking)이라고도 한다. 이때, 주파수 대역에 따라 마스킹되는 음의 임계값도 달라진다.

<14> 심리 음향 모델을 사용하여 필터 뱅크의 각 서브밴드에서 들을 수 없는 최대 잡음 모델을 결정할 수 있는데, 이 각각의 서브 밴드에서의 잡음 레벨, 즉 마스킹 임계값을 사용하여, 각 서브 밴드에 대한 SMR 값을 구하는 것이 가능하다.

<15> 심리 음향 모델을 사용한 부호화 방법은 양수인이 Motorola, Inc이고, 발명의 명칭이 "System and method of encoding and decoding a layered bitstream by re-applying psychoacoustic analysis in the decoder"인, 미국 특허 제6,092,041호에 개시되어 있다.

<16> 도 1은 일반적인 MPEG 오디오 부호화기를 도시하는 도면이다. 여기에서는, MPEG 오디오 부호화기 중 MPEG-1 레이어 3, 즉 MP 3 오디오 부호화기를 예를 들어 설명한다.

<17> MP 3 오디오 부호화기는 필터 뱅크(filter bank)(110), 변형 이산 여현 변환부(modified discrete cosine transform: MDCT)(120), 고속 푸리에 변환부(fast fourier transform: FFT)(130), 심리 음향 부호화부(psychoacoustic model unit)(140), 양자화 및 허프만 인코딩부(150), 비트 스트림 포매팅부(160)를 포함한다.

- <18> 필터 뱅크(110)는 오디오 신호의 통계적인 중복성을 제거하기 위해 입력된 시간 영역의 오디오 신호를 32개의 주파수 영역의 서브 밴드로 세분한다.
- <19> MDCT부(120)는 주파수 분해능(frequency resolution)을 증가시키기 위해, 심리 음향 모델부(140)로부터 입력된 윈도우 스위칭 정보를 이용하여 필터 뱅크(110)에서 분할된 서브 밴드를 보다 세밀한 주파수 대역으로 분할한다. 예를 들어, 심리 음향 모델부(140)로부터 입력된 윈도우 스위칭 정보가 롱 윈도우(long window)를 표시하는 경우에는, 36 포인트 MDCT를 사용하여 32개의 서브 밴드보다 더 세밀하게 주파수 대역을 분할하고, 윈도우 스위칭 정보가 쇼트 윈도우(short window)를 표시하는 경우에는, 12 포인트 MDCT를 사용하여 32개의 서브 밴드보다 더 세밀하게 주파수 대역을 분할한다.
- <20> FFT부(130)는 입력된 오디오 신호를 주파수 영역의 스펙트럼으로 변환하여 심리 음향 모델부(140)로 출력한다.
- <21> 심리 음향 모델부(140)는 인간의 청각 특성에 의한 지각적인 중복성을 제거하기 위해, FFT부(130)에서 출력된 주파수 스펙트럼을 이용하여, 각각의 서브 밴드에 대한 귀에 들리지 않는 잡음 레벨인 마스킹 임계값 (masking threshold), 즉 신호 대 마스크율 (signal to mask ratio: SMR)을 결정한다. 심리 음향 모델부(140)에서 결정된 SMR 값은 양자화 및 허프만 부호화부(120)로 입력된다.
- <22> 또한, 심리 음향 모델부(140)는 지각 에너지(perceptual energy)를 계산하여 윈도우 스위칭 여부를 결정하여, 윈도우 스위칭 정보를 MDCT부(120)로 출력한다.
- <23> 양자화 및 허프만 부호화부(150)에서는 심리 음향 모델부(140)에서 입력된 SMR 값에 기초하여, MDCT부(120)로부터 입력된 MDCT가 수행된 주파수 영역의 데이터에 대해,

지각적 중복성을 제거하기 위한 비트 할당과 오디오 부호화를 위한 양자화 과정을 수행한다.

<24> 비트 스트림 포매팅부(160)는 양자화 및 허프만 부호화부(150)로부터 입력된 부호화된 오디오 신호를 MPEG에서 정한 비트 스트림으로 포매팅하여 출력한다.

<25> 상기에서 설명된 바와 같이, 도 1에 도시된 종래의 심리 음향 모델에서는 마스킹 임계치를 계산하기 위해 입력 오디오 신호로부터 얻어진 FFT 스펙트럼을 사용한다. 하지만, 필터 뱅크는 앨리어싱(aliasing)을 일으키고, 이들 앨리어싱이 일어난 성분들로부터 얻어진 값들이 양자화 단계에서 사용되기 때문에, 심리 음향 모델에서 FFT 스펙트럼에 기초하여 SMR을 구하고, 이를 양자화 단계에서 사용하는 경우 최적의 결과를 얻을 수 없다는 문제점이 있다.

【발명이 이루고자 하는 기술적 과제】

<26> 본 발명은 상기와 같은 문제점을 해결하기 위한 것으로서, 변형된 심리 음향 모델을 사용하여 종래의 MPEG 오디오 부호화기에 비해 출력 오디오 스트림의 음질을 향상시키고, 디지털 오디오 부호화 단계의 계산량을 감소시키는 것이 가능한 디지털 오디오 부호화 방법 및 장치를 제공하는 것을 목적으로 한다.

【발명의 구성 및 작용】

<27> 상기의 기술적 과제를 해결하기 위하여, 본 발명에 따른 디지털 오디오 부호화 방법은 입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 단계와, 상기 결정된 윈도우 타입에 따라 상기 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 단계와, 상기 결정된 윈도우 타입을 이용하여 상기 입력 오디오 신호로부터 FFT 스펙트럼을 생성하

는 단계와, 상기 생성된 CMDCT 스펙트럼 및 FFT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 단계를 포함하는 것을 특징으로 한다.

<28> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 방법은 결정된 윈도우 타입이 롱 윈도우인 경우, 롱 윈도우를 적용하여 롱 CMDCT 스펙트럼을 생성하고, 쇼트 윈도우를 적용하여 쇼트 FFT 스펙트럼을 생성하고, 생성된 롱 CMDCT 스펙트럼 및 쇼트 FFT 스펙트럼에 기초하여 심리 음향 모델 분석을 수행하는 것을 특징으로 한다.

<29> 상기의 기술적 과제를 해결하기 위하여, 본 발명에 따른 디지털 오디오 부호화 장치는 입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 윈도우 스위칭부와, 상기 윈도우 스위칭부에서 결정된 윈도우 타입에 따라 상기 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 CMDCT부와, 상기 윈도우 스위칭부에서 결정된 윈도우 타입을 이용하여 상기 입력 오디오 신호로부터 FFT 스펙트럼을 생성하는 FFT부와, 상기 CMDCT부에서 생성된 CMDCT 스펙트럼 및 상기 FFT부에서 생성된 FFT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 심리 음향 모델부를 포함하는 것을 특징으로 한다.

<30> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 장치는 윈도우 스위칭부에서 결정된 윈도우 타입이 롱 윈도우인 경우, 상기 CMDCT부는 롱 윈도우를 적용하여 롱 CMDCT 스펙트럼을 생성하고, 상기 FFT부는 쇼트 윈도우를 적용하여 쇼트 FFT 스펙트럼을 생성하고, 상기 심리 음향 모델부는 상기 CMDCT부에서 생성된 롱 CMDCT 스펙트럼 및 상기 FFT부에서 생성된 쇼트 FFT 스펙트럼에 기초하여 심리 음향 모델 분석을 수행하는 것을 특징으로 한다.

- <31> 상기의 기술적 과제를 해결하기 위하여, 본 발명에 따른 디지털 오디오 부호화 방법은 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 단계와, 생성된 CMDCT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 단계를 포함하는 것을 특징으로 한다.
- <32> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 방법은 입력 오디오 신호에 대해 롱 윈도우 및 쇼트 윈도우를 적용하여 CMDCT를 수행하여, 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 생성하는 단계를 더 포함하는 것을 특징으로 한다.
- <33> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 방법은 생성된 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 사용하여 심리 음향 모델 분석을 수행하는 것을 특징으로 한다.
- <34> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 방법은 결정된 윈도우 타입이 롱 윈도우인 경우에는, 심리 음향 모델 분석 결과에 기초하여 롱 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하고, 결정된 윈도우 타입이 쇼트 윈도우인 경우에는, 심리 음향 모델 분석 결과에 기초하여 쇼트 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하는 것을 특징으로 한다.
- <35> 상기의 기술적 과제를 해결하기 위하여, 본 발명에 따른 디지털 오디오 부호화 장치는 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 CMDCT부와, 상기 CMDCT부에서 생성된 CMDCT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 심리 음향 모델부를 포함하는 것을 특징으로 한다.

- <36> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 장치에서는 상기 CMDCT부는 상기 입력 오디오 신호에 대해 롱 윈도우 및 쇼트 윈도우를 적용하여 CMDCT를 수행하여, 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 생성하는 것을 특징으로 한다.
- <37> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 장치에서는 상기 심리 음향 모델부는 상기 CMDCT부에서 생성된 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 사용하여 심리 음향 모델 분석을 수행하는 것을 특징으로 한다.
- <38> 상기 기술적 과제를 달성하기 위한, 본 발명에 따른 보다 바람직한 디지털 오디오 부호화 장치는 양자화 및 부호화부를 더 포함하며, 양자화 및 부호화부는 결정된 윈도우 타입이 롱 윈도우인 경우에는, 심리 음향 모델 분석 결과에 기초하여 롱 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하고, 결정된 윈도우 타입이 쇼트 윈도우인 경우에는, 심리 음향 모델 분석 결과에 기초하여 쇼트 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하는 것을 특징으로 한다.
- <39> MPEG 오디오 부호화기는 매우 많은 계산량을 요구하기 때문에, 실시간 처리에 적용하는 것이 곤란하다. 출력 오디오의 음질을 저하시킴으로써 인코딩 알고리즘을 단순화하는 것이 가능하다. 하지만, 음질을 저하시키지 않고서 계산량을 감소시키는 것은 아주 어려운 일이다.
- <40> 또한, 종래의 MPEG 오디오 부호화기에서 사용되는 필터-뱅크는 앨리어싱(aliasing)을 일으킨다. 이들 앨리어싱이 일어난 성분들로부터 얻어진 값들이 양자화 단계에서 사

용되기 때문에, 이들 앨리어싱이 일어난 스펙트럼에 심리 음향 모델을 적용하는 것이 바람직하다.

<41> 또한, 후술하는 수학적식 2에서 보여지는 바와 같이, MDCT 스펙트럼은 주파수 $2\pi(k+0.5)/N$, $k=0,1, \dots, N/2 - 1$ 에서의 크기 및 위상 값을 준다. 따라서, 이들 주파수에서의 스펙트럼을 계산하여, 심리 음향 모델을 적용하는 것이 바람직하다.

<42> 또한, 필터-뱅크의 출력에 대해 CMDCT를 적용하여 입력 신호의 스펙트럼을 계산하고, 이에 따라 심리 음향 모델을 적용함으로써, 종래의 MPEG 오디오 부호화기에 비해 FFT 변환에 필요한 계산량을 줄이거나, FFT 변환 과정을 생략하는 것이 가능하다.

<43> 본 발명은 상기와 같은 점에 착안한 것으로서, 본 발명에 따른 오디오 부호화 방법 및 장치는 출력되는 MPEG 오디오 스트림의 음질을 저하시키지 않으면서, MPEG 오디오 부호화 프로세서의 복잡도를 감소시키는 것이 가능하다.

<44> 이하에서는, 수학적식 1 내지 4를 참조하여, 본 발명에 사용되는 알고리즘을 상세히 설명한다.

<45> 필터-뱅크는 입력 신호를 $\pi/32$ 의 해상도로 입력 신호를 분할한다. 아래에서 설명되는 바와 같이, 필터-뱅크의 출력 값에 CMDCT를 적용함으로써 입력 신호의 스펙트럼을 계산하는 것이 가능하다. 이때 변환 길이(transform length)는 필터 뱅크의 출력 값을 사용하지 않고, 입력 신호에 CMDCT를 직접 적용한 경우보다 훨씬 짧다. 필터-뱅크 출력에 이러한 짧은 길이의 변환 값을 사용하는 것은 긴 길이의 변환 값을 사용하는 경우보다 계산량을 줄일 수 있다는 장점이 있다.

<46> CMDCT는 아래 수학적식 1에 의해 계산될 수 있다.

<47>

【수학식 1】
$$\tilde{X}(k) = X_c(k) + jX_s(k)$$

<48>

여기에서, $k = 0, 1, 2, \dots, N/2 - 1$ 이다.

<49>

이 경우, $X_c(k)$ 는 MDCT(modified discrete cosine transform)이고, $X_s(k)$ 는 MDST(modified discrete sine transform)이다. 아래 유도식들은 CMDCT와 FFT 간의 관계를 설명한다.

<50>

$$X_c(k) = \sum_{n=0}^{N-1} x(n) \cos\{2\pi(k+0.5)(n+0.5+N/4)/N\}$$

【수학식 2】
$$= \sum_{n=0}^{N-1} x(n) \cos\{2\pi n(k+0.5)/N + \Phi_k\}$$

<51>

여기에서, $\Phi_k = 2\pi(k+0.5)(N/4+0.5)/N$ 이고, $k = 0, 1, \dots, N/2 - 1$ 이다. 또한, MDST는 MDCT와 동일하게,

<52>

$$X_s(k) = \sum_{n=0}^{N-1} x(n) \sin\{2\pi(k+0.5)(n+0.5+N/4)/N\}$$

$$= \sum_{n=0}^{N-1} x(n) \sin\{2\pi n(k+0.5)/N + \Phi_k\}$$

【수학식 3】

<53>

여기에서, $k = 0, 1, \dots, N/2 - 1$ 이다.

<54>

또한, 아래 수학식 4와 같이, $\bar{x}(k)$ 를 CMDCT의 공액 복소수(complex conjugate)라고 두면,

<55>

$$\begin{aligned}\bar{x}(k) &= X_c(k) - jX_s(k) \\ &= \sum_{n=0}^{N-1} e^{j\{2\pi n(k+0.5)/N + \phi_k\}} \\ &= e^{-j\phi_k} X'(k)\end{aligned}$$

【수학식 4】

<56>

$$X'(k) = \sum_{n=0}^{N-1} e^{j\{2\pi n(k+0.5)/N\}}$$

여기에서, 이고, $k = 0, 1, 2, \dots, N/2 - 1$ 이다.

<57>

상기 수학식 4에서 알 수 있듯이, CMDCT의 공액 복소수는 DFT 스펙트럼의 주파수 사이, 즉 $2\pi(K+0.5)/N$, $k=0, 1, \dots, N/2 - 1$ 의 주파수들에서 스펙트럼을 계산한다.

<58>

CMDCT의 위상은 $X'(k)$ 의 위상이 쉬프트된 것이고, 이러한 위상 쉬프트는 MPEG-1 레이어 3의 심리 음향 모델에서의 비예측도(unpredictability measure) 계산에 영향을 미치지 않는다.

<59>

본 발명에 따른 심리 음향 모델에서는 이러한 점을 고려하여, 심리 음향 모델 분석 수행시 FFT 스펙트럼 대신 CMDCT 스펙트럼을 사용하거나, 또는 롱 FFT 스펙트럼 또는 쇼트 FFT 스펙트럼 대신 롱 CMDCT 스펙트럼 또는 CMDCT 스펙트럼을 사용한다. 이에 따라, FFT 변환에 소요되는 계산량을 감소시키는 것이 가능하다.

<60>

이하에서는, 실시예들에 기초하여 본 발명을 상세히 설명한다.

<61>

도 2는 본 발명의 일 실시예에 따른 오디오 부호화 장치를 도시하는 블록도이다.

<62>

필터 बैं크(210)는 입력 오디오 신호의 통계적인 중복성을 제거하기 위해 입력된 시간 영역의 오디오 신호를 주파수 영역의 서브 밴드들로 분할한다. 본 실시예에서는 $\pi/32$ 의 대역폭을 갖는 32개의 서브 밴드로 분할한다. 본 실시예에서는 32 폴리 페이즈

필터 बैं크(poly phase filter bank)를 사용하였지만, 선택적으로 서브 밴드 부호화가 가능한 다른 필터를 사용하는 것도 가능하다.

<63> 윈도우 스위칭부(window switching unit)(220)는 입력 오디오 신호의 특성에 기초하여 CMDCT부(230) 및 FFT부(240)에서 사용될 윈도우 타입을 결정하고, 결정된 윈도우 타입에 대한 정보를 CMDCT부(230) 및 FFT부(240)로 입력한다.

<64> 윈도우 타입에는 쇼트 윈도우(short window)와 롱 윈도우(long window)가 있다. MPEG-1 레이어 3에서는 롱 윈도우, 스타트 윈도우(start window), 쇼트 윈도우, 스톱 윈도우(stop window) 등을 규정하고 있다. 이때, 스타트 윈도우 또는 스톱 윈도우는 롱 윈도우에서 쇼트 윈도우로 스위칭하기 위해 사용된다. 본 실시예에서는 MPEG-1에 규정된 윈도우 타입을 예를 들어 설명하였지만, 선택적으로 다른 윈도우 타입에 따라 윈도우 스위칭 알고리즘을 수행하는 것도 가능하다. 본 발명에 따른 윈도우 스위칭 알고리즘에 대한 상세한 설명은 도 3 및 4를 참조하여 후술한다.

<65> CMDCT(complex modified discrete cosine transform)부(230)는 윈도우 스위칭부(220)으로부터 입력된 윈도우 타입 정보에 기초하여 필터 बैं크(210)의 출력 데이터에 롱 윈도우 또는 쇼트 윈도우를 적용하여 CMDCT를 수행한다.

<66> CMDCT부(230)에서 계산된 CMDCT의 실수 값, 즉 MDCT 값은 양자화 및 부호화부(260)로 입력된다. 또한, CMDCT부(230)에서는 계산된 서브-밴드 스펙트럼을 합쳐서 전체 스펙트럼(full spectrum)을 계산하고, 계산된 전체 스펙트럼을 심리 음향 모델부(250)로 전송한다. 서브-밴드 스펙트럼으로부터 전체 스펙트럼을 구하는 과정은 도 5와 관련하여 후술한다.

- <67> 선택적으로, MDCT의 빠른 실행을 위해 LAME 알고리즘이 사용될 수 있다. LAME 알고리즘에서, MDCT는 아래 수학적 식 1을 전개(unrolling)함으로써 최적화된다. 계산에 관련된 삼각법(trigonometric)에 의한 계수들의 대칭성을 이용함으로써, 동일한 계수들에 의한 연속하는 승산 연산은 가산 연산으로 대체된다. 이는, 244번의 승산 및 324번의 가산으로 연산 카운트를 감소시키고, 36 포인트 MDCT에 대해 약 70% 정도의 MDCT 시간을 절감한다. 이 알고리즘은 MDST에 대해서도 또한 적용될 수 있다.
- <68> FFT부(240)는 윈도우 스위칭부(220)로부터의 윈도우 타입 정보에 기초하여 입력 오디오 신호에 대해 롱 윈도우 또는 쇼트 윈도우를 사용하여, FFT를 수행하고, 계산된 롱 FFT 스펙트럼 또는 쇼트 FFT 스펙트럼을 심리 음향 모델부(250)로 출력한다. 이때, CMDCT부(230)에서 사용되는 윈도우 타입이 롱 윈도우인 경우에는 FFT부(240)에서는 쇼트 윈도우를 사용한다. 즉, CMDCT부(230)의 출력이 롱 CMDCT 스펙트럼인 경우, FFT부(240)의 출력은 쇼트 FFT 스펙트럼이 된다. 마찬가지로, CMDCT부(230)의 출력이 쇼트 CMDCT 스펙트럼인 경우, FFT부(240)의 출력은 롱 FFT 스펙트럼이 된다.
- <69> 심리 음향 모델부(250)는 CMDCT부(230)로부터의 CMDCT 스펙트럼 및 FFT부(240)로부터의 FFT 스펙트럼을 조합하여 심리 음향 모델에서 사용되는 비예측도(unpredictability)를 계산한다.
- <70> 예를 들어, CMDCT에서 롱 윈도우가 사용되는 경우, 롱 스펙트럼은 롱 MDCT와 롱 MDST의 결과 값을 사용하여 계산되며, 쇼트 스펙트럼은 FFT를 사용하여 계산된다. 여기에서, 롱 스펙트럼의 경우, CMDCT부(230)에서 계산된 CMDCT 스펙트럼을 사용하는 것은, 수학적 식 3 및 수학적 식 4로부터 알 수 있듯이 FFT 및 MDCT의 크기는 유사하다는 점을 이용한 것이다.

- <71> 또한, CMDCT에서 쇼트 윈도우가 사용되는 경우, 쇼트 스펙트럼은 쇼트 MDCT와 쇼트 MDST의 결과 값을 사용하여 계산되며, 롱 스펙트럼은 FFT 스펙트럼을 사용하여 계산된다.
- <72> 한편, CMDCT부(230)에서 계산된 CMDCT 스펙트럼은 롱 윈도우가 적용된 경우에는 1152 (32 서브 밴드 \times 36 서브-서브 밴드)의 길이, 쇼트 윈도우가 적용된 경우에는 384 (32 서브 밴드 \times 12 서브-서브 밴드)의 길이를 갖는다. 반면에, 심리 음향 모델부(250)는 길이가 1024 또는 256인 스펙트럼을 필요로 한다.
- <73> 따라서, CMDCT 스펙트럼은 심리 음향 모델 분석이 수행되기 전에 선형 매핑(linear mapping)에 의해 1152 (또는 384)의 길이에서 1024 (또는 256)의 길이로 재-샘플링된다.
- <74> 또한, 심리 음향 모델부(250)에서는 계산된 비예측도를 사용하여, SMR 값을 구하고, 이를 양자화 및 부호화부(260)로 출력한다.
- <75> 양자화 및 부호화부(260)는 스케일 팩터(scale factor)를 결정하고, 심리 음향 모델부(250)에서 계산된 SMR 값에 기초하여 양자화 계수를 결정한다. 결정된 양자화 계수에 기초하여 양자화를 수행하고, 양자화된 데이터에 대해 허프만 부호화를 수행한다.
- <76> 비트 스트림 포맷 부(270)는 양자화 및 부호화부(260)로부터 입력된 데이터를 특정 포맷으로 변환하여 출력한다. 상기 오디오 부호화 장치가 MPEG 오디오 부호화 장치인 경우에는 MPEG 표준에서 정한 포맷으로 변환하여 출력한다.
- <77> 도 3은 도 2의 윈도우 스위칭부(220)에서 사용되는 필터 뱅크 출력에 기초한 윈도우 스위칭 알고리즘에 사용되는 천이 신호 검출(transient signal detection) 방식을 도시하는 도면이다.

- <78> MPEG에서 표준화된 MPEG 오디오 규격에 따르면, 실제 윈도우 타입은 현재 프레임의 윈도우 타입과 다음 프레임의 윈도우-스위칭 플래그(window-switching flag)에 기초하여 결정된다. 심리 음향 모델은 지각 엔트로피(perceptual entropy)에 기초하여 윈도우 스위칭 플래그를 결정한다. 때문에, 심리 음향 모델은 필터-뱅크 및 MDCT에서 처리되는 프레임 보다 적어도 하나의 이전 프레임에 대해 수행하는 것이 필요했다.
- <79> 반면에, 본 발명에 따른 심리 음향 모델은 상기에서 설명한 바와 같이 CMDCT 스펙트럼을 사용한다. 따라서, 윈도우 타입은 CMDCT가 적용되기 전에 결정되어야 한다. 또한, 이러한 이유로, 윈도우-스위칭 플래그는 필터-뱅크 출력으로부터 결정되고, 필터-뱅크 및 윈도우-스위칭은 양자화 및 심리 음향 모델에 비해 한 프레임 앞선 프레임에 대해 수행된다.
- <80> 도 3에 도시된 바와 같이, 필터 뱅크로부터의 입력 신호는 3개의 시간 대역과 2개의 주파수 대역, 즉 총 6개의 대역으로 분할된다. 도 3에서, 가로축은 각 프레임당 36개의 샘플, 즉 각각 12개의 샘플을 갖는 3개의 시간 대역으로 나눈다. 세로축은 각 프레임당 32개의 서브-밴드, 즉 각각 16개의 서브 밴드를 갖는 2개의 주파수 대역으로 나눈다. 여기에서, 36 개의 샘플과 32개의 서브-밴드는 1152개의 샘플 입력에 대응한다.
- <81> 빗금친 부분은 천이 검출을 위해 사용되는 부분인데, 설명의 편의를 위해 각 빗금친 부분을 (1), (2), (3), 및 (4)라고 한다. 각 영역에 대한 에너지가 E1, E2, E3, 및 E4라고 하는 경우, 영역 (1) 및 (2) 간의 에너지 비 E1/E2와 영역 (3) 및 (4) 간의 에너지 비 E3/E4는 천이 여부를 표시하는 천이 표시자(transient indicator)이다.

- <82> 비-천이 신호(non_transient signal)의 경우, 천이 표시자의 값은 일정 범위내에 있다. 따라서, 천이 표시자가 일정 범위를 벗어나는 경우, 윈도우 스위칭 알고리즘은 쇼트 윈도우가 필요하다는 것을 표시한다.
- <83> 도 4는 도 2에 도시된 윈도우 스위칭부(220)에서 수행되는 윈도우 스위칭 알고리즘 방식을 도시하는 플로우 차트이다.
- <84> 단계 410에서는 32개의 서브 밴드와, 각 서브 밴드당 36개의 출력 샘플을 갖는 한 프레임의 필터-뱅크 출력이 입력된다.
- <85> 단계 420에서는 도 3에 도시된 바와 같이, 각각 12개의 샘플 값을 갖는 3개의 시간 대역과 16개의 주파수 대역을 갖는 주파수 대역으로 분할된다.
- <86> 단계 430에서는 천이 신호를 검출하기 위해 사용되는 밴드들의 에너지 E1, E2, E3, 및 E4가 계산된다.
- <87> 단계 440에서는 입력 신호에 천이(transient)가 있는지 여부를 판단하기 위해 단계 430에서 계산된 주변 밴드들의 에너지들이 비교된다. 즉, E1/E2 및 E3/E4가 계산된다.
- <88> 단계 450에서는 계산된 주변 밴드들의 에너지 비에 기초하여 입력 신호에 천이가 있는지 여부를 결정한다. 입력 신호에 천이가 있는 경우에는 쇼트 윈도우를 표시하기 위한 윈도우 스위칭 플래그가 생성되고, 천이가 없는 경우에는 롱 윈도우를 표시하기 위한 윈도우 스위칭 플래그가 생성된다.
- <89> 단계 460에서는 단계 450에서 생성된 윈도우 스위칭 플래그와 이전 프레임에서 사용된 윈도우에 기초하여 실제 적용되는 윈도우 타입을 결정한다. 적용되는 윈도우 타입

들은 MPEG-1 표준에서 사용되고 있는 "쇼트", "롱 스톱", "롱 스타트", 또는 "롱" 중 하나일 수 있다.

<90> 도 5는 본 발명에 따른 서브-밴드 스펙트럼(sub-band spectra)로부터 전체 스펙트럼을 구하는 방법을 도시하는 도면이다.

<91> 이하에서는, 도 5를 참조하여 서브-밴드 필터-뱅크의 출력으로부터 계산된 스펙트럼으로부터 신호 스펙트럼을 근사적으로 계산하기 위한 방법을 설명한다.

<92> 도 5에 도시된 바와 같이 입력 신호는 분석 필터들(analysis filters), $H_0(Z)$, $H_1(Z)$, $H_2(Z)$, . . . $H_{M-1}(Z)$ 에 의해 필터링되고, 다운샘플링된다. 이후, 다운샘플링되었던 신호, $y_0(n)$, $y_1(n)$, $y_2(n)$, . . . $y_{M-1}(n)$ 은 업샘플링되고(upsampled), 합성 필터들(synthesis filters), $G_0(Z)$, $G_1(Z)$, $G_2(Z)$, . . . $G_{M-1}(Z)$ 에 의해 필터링되고, 신호를 재구성하기 위해 합쳐진다.

<93> 이러한 과정은, 주파수 영역에서의, 스펙트럼을 반복하고, 대응하는 필터의 주파수 응답으로 승산한 후, 모든 대역의 스펙트럼을 합치는 과정에 대응한다. 따라서, 이 필터들이 이상적인 경우, 각각의 대역에 대한 $Y_m(k)$ 를 모두 합친 스펙트럼과 같게되고, 결과적으로 입력 FFT 스펙트럼을 얻을 수 있다. 또한, 이들 필터들이 이상적인 필터에 근접한 경우에도, 근사적인 스펙트럼을 얻을 수 있는데, 본 발명에 따른 심리 음향 모델에서는 이를 이용한다.

<94> 실험 결과, 사용되는 필터들이 이상적인 밴드-패스 필터(band-pass filter)가 아닌 경우에도, MPEG-1 레이어 3에 사용되는 필터-뱅크인 경우, 상기 방법에 의해 얻어진 스펙트럼은 실제 스펙트럼과 유사하다는 실험 결과를 얻었다.

- <95> 이와 같이, 입력 신호의 스펙트럼은 모든 대역에서의 CMDCT 스펙트럼들을 합침으로써 얻을 수 있다. CMDCT를 사용하여 얻어진 스펙트럼은 1152 포인트인 반면, 심리 음향 모델에 필요한 스펙트럼은 1024 포인트이다. 따라서, CMDCT 스펙트럼은 간단한 선형 매핑을 사용하여 재 샘플링된 후 심리 음향 모델에서 사용될 수 있다.
- <96> 도 6은 본 발명의 또 다른 실시예에 따른 오디오 부호화 방법을 도시하는 플로우 차트이다.
- <97> 단계 610에서는 필터 बैं크에서 오디오 신호를 입력 받고, 입력된 오디오 신호의 통계적인 중복성을 제거하기 위해 입력된 시간 영역의 오디오 신호를 주파수 영역의 서브 밴드들로 분할한다.
- <98> 단계 620에서는 입력 오디오 신호의 특성에 기초하여 윈도우 타입을 결정한다. 입력 신호가 천이 신호인 경우에는 단계 630으로 진행하고, 입력 신호가 천이 신호가 아닌 경우에는 단계 640으로 진행한다.
- <99> 단계 630에서는 단계 610에서 처리된 오디오 데이터에 대해, 쇼트 윈도우를 적용하여 쇼트 CMDCT를 수행하고, 또한 동시에 롱 윈도우를 적용하여 롱 FFT를 수행한다. 이 결과, 쇼트 CMDCT 스펙트럼 및 롱 FFT 스펙트럼을 얻는다.
- <100> 단계 640에서는 단계 610에서 처리된 오디오 데이터에 대해, 롱 윈도우를 적용하여 롱 CMDCT를 수행하고, 또한 동시에 쇼트 윈도우를 적용하여 쇼트 FFT를 수행한다. 이 결과, 롱 CMDCT 스펙트럼 및 쇼트 FFT 스펙트럼을 얻는다.
- <101> 단계 650에서는 단계 620에서 결정된 윈도우 타입이 쇼트 윈도우인 경우에는 단계 630에서 얻어진 쇼트 CMDCT 스펙트럼 및 롱 FFT 스펙트럼을 이용하여 심리 음향 모델에

서 사용되는 비예측도를 계산하고, 단계 620에서 결정된 윈도우 타입이 롱 윈도우인 경우에는 단계 640에서 얻어진 롱 CMDCT 스펙트럼 및 쇼트 FFT 스펙트럼을 이용하여 비예측도를 계산한다. 또한, 계산된 비예측도에 기초하여 SMR 값을 계산한다.

<102> 단계 660에서는 단계 610에서 얻어진 오디오 데이터에 대해 단계650에서 계산된 SMR 값에 따라 양자화를 수행하고, 양자화된 데이터에 대해 허프만 부호화를 수행한다.

<103> 단계 670에서는 단계660에서 부호화된 데이터를 특정 포맷으로 변환하여 출력한다. 상기 오디오 부호화 방법이 MPEG 오디오 부호화 방법인 경우에는 MPEG 표준에서 정한 포맷으로 변환하여 출력한다.

<104> 도 7은 본 발명의 또 다른 실시예에 따른 오디오 부호화기를 설명하는 도면이다.

<105> 도 7에 도시된 오디오 부호화기는 필터 뱅크부(710), 윈도우 스위칭부(720), CMDCT부(730), 심리 음향 모델부(740), 양자화 및 부호화부(750), 및 비트 스트림 포매팅부(760)로 이루어진다.

<106> 여기에서, 필터 뱅크부(710), 양자화 및 부호화부(750), 및 비트 스트림 포매팅부(760)는 도 2의 필터 뱅크부(210), 양자화 및 부호화부(260), 및 비트 스트림 포매팅부(270)와 유사한 기능을 수행하므로, 설명의 간단을 위해 상세한 설명은 생략한다.

<107> 윈도우 스위칭부(720)는 입력 오디오 신호의 특성에 기초하여 CMDCT부(730)에서 사용될 윈도우 타입을 결정하고, 결정된 윈도우 타입 정보를 CMDCT부(730)로 전송한다.

<108> CMDCT부(730)는 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 함께 계산한다. 본 실시예에서는, 심리 음향 모델부(740)에서 사용되는 롱 CMDCT 스펙트럼은 36 포인트 CMDCT를 수행하고, 이를 모두 합친 후, 1152 길이의 스펙트럼을 1024 길이의 스펙트럼으

로 재-샘플링함으로써 얻어진다. 또한, 심리 음향 모델부(740)에서 사용되는 쇼트 CMDCT 스펙트럼은 12 포인트 CMDCT를 수행하고, 이를 모두 합친 후, 그 결과인 384 길이의 스펙트럼을 256 길이의 스펙트럼으로 재-샘플링함으로써 얻어진다.

<109> CMDCT부(730)는 계산된 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 심리 음향 모델부(740)로 출력한다. 또한, CMDCT부(730)는 윈도우 스위칭부(720)로부터 입력된 윈도우 타입이 롱 윈도우 인 경우에는, 롱 MDCT 스펙트럼을 양자화 및 부호화부(750)로 입력하고, 입력된 윈도우 타입이 쇼트 윈도우 인 경우에는 쇼트 MDCT 스펙트럼을 양자화 및 부호화부(750)로 입력한다.

<110> 심리 음향 모델부(740)는 CMDCT부(730)로부터 전송된 롱 스펙트럼 및 쇼트 스펙트럼에 따라 비예측도를 계산하고, 계산된 비예측도에 기초하여 SMR 값을 계산하여, 양자화 및 부호화부(750)로 전송한다.

<111> 양자화 및 부호화부(750)는 CMDCT부(730)로부터 전송된 롱 MDCT 스펙트럼 및 쇼트 MDCT 스펙트럼과, 심리 음향 모델부로부터 입력된 SMR 정보에 기초하여, 스케일 팩터 및 양자화 계수를 결정한다. 결정된 양자화 계수에 기초하여 양자화를 수행하고, 양자화된 데이터에 대해 허프만 부호화를 수행한다.

<112> 비트 스트림 포맷 부(760)는 양자화 및 부호화부(750)로부터 입력된 데이터를 특정 포맷으로 변환하여 출력한다. 상기 오디오 부호화 장치가 MPEG 오디오 부호화 장치인 경우에는 MPEG 표준에서 정한 포맷으로 변환하여 출력한다.

<113> 도 8은 본 발명의 또 다른 실시예에 따른 오디오 부호화 방법을 도시하는 플로우 차트이다.

- <114> 단계 810에서는 필터 बैं크에서 오디오 신호를 입력 받고, 입력된 오디오 신호의 통계적인 중복성을 제거하기 위해 입력된 시간 영역의 오디오 신호를 주파수 영역의 서브 밴드들로 분할한다.
- <115> 단계 820에서는 입력 오디오 신호의 특성에 기초하여 윈도우 타입을 결정한다.
- <116> 단계 830에서는 단계 810에서 처리된 오디오 데이터에 대해, 쇼트 윈도우를 적용하여 쇼트 CMDCT를 수행하고, 또한 동시에 롱 윈도우를 적용하여 롱 CMDCT를 수행한다. 이 결과, 쇼트 CMDCT 스펙트럼 및 롱 CMDCT 스펙트럼을 얻는다.
- <117> 단계 840에서는 단계 830에서 얻어진 쇼트 CMDCT 스펙트럼 및 롱 CMDCT 스펙트럼을 이용하여 심리 음향 모델에서 사용되는 비예측도를 계산한다. 또한, 계산된 비예측도에 기초하여 SMR 값을 계산한다.
- <118> 단계 850에서는 단계 820에서 결정된 윈도우 타입이 롱 윈도우인 경우에는 단계 830에서 얻어진 스펙트럼 중 롱 MDCT 값을 입력받아, 이에 대해 단계 840에서 계산된 SMR 값에 따라 양자화를 수행하고, 양자화된 데이터에 대해 허프만 부호화를 수행한다.
- <119> 단계 860에서는 단계 850에서 부호화된 데이터를 특정 포맷으로 변환하여 출력한다. 상기 오디오 부호화 장치가 MPEG 오디오 부호화 장치인 경우에는 MPEG 표준에서 정한 포맷으로 변환하여 출력한다.
- <120> 본 발명은 상술한 실시예에 한정되지 않으며, 본 발명의 사상내에서 당업자에 의한 변형이 가능함은 물론이다. 특히, 본 발명은 MPEG-1 레이어 3 뿐만 아니라 MDCT 및 심리 음향 모델을 사용하는 MPEG-2 AAC, MPEG 4, WMA 등과 같은 모든 오디오 부호화 장치 및 방법에 적용될 수 있다.

<121> 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 하드디스크, 플로피디스크, 플래쉬 메모리, 광데이터 저장장치 등이 있으며, 또한 캐리어 웨이브 (예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드로서 저장되고 실행될 수 있다.

【발명의 효과】

<122> 상술한 바와 같이 본 발명에 따른 개선된 심리 음향 모델을 적용하여, FFT 스펙트럼 대신 CMDCT 스펙트럼을 사용함으로써, 입력 오디오 신호에 비해 출력 오디오 스트림의 음질의 저하 없이 FFT 변환에 소요되는 계산량 및 MPEG 오디오 부호화기의 복잡도를 감소시키는 것이 가능하다는 효과가 있다.

【특허청구범위】**【청구항 1】**

디지털 오디오 부호화 방법에 있어서,

(a) 입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 단계와,

(b) 상기 결정된 윈도우 타입에 따라 상기 입력 오디오 신호로부터 CMDCT(complex modified discrete cosine transform) 스펙트럼을 생성하는 단계와,

(c) 상기 결정된 윈도우 타입을 이용하여 상기 입력 오디오 신호로부터 FFT(fast fourier transform) 스펙트럼을 생성하는 단계와,

(d) 상기 생성된 CMDCT 스펙트럼 및 FFT 스펙트럼을 이용하여 심리 음향 모델 (psychoacoustic model) 분석을 수행하는 단계를 포함하는 것을 특징으로 하는 방법.

【청구항 2】

제1항에 있어서, 상기 (a) 단계는 (a1) 상기 입력 오디오 신호를 필터링하여 복수 개의 서브 밴드로 분할하는 단계를 더 포함하며, 상기 윈도우 타입을 결정하는 단계는 상기 서브 밴드로 분할된 데이터에 대해 수행되는 것을 특징으로 하는 방법.

【청구항 3】

제2항에 있어서, 상기 (a1) 단계는 다상 필터 뱅크(polyphase filter bank)에 의해 수행되는 것을 특징으로 하는 방법.

【청구항 4】

제1항에 있어서, 상기 (a) 단계에서 결정된 윈도우 타입이 롱 윈도우(long window) 인 경우, 상기 (b) 단계에서는 롱 윈도우를 적용하여 롱 CMDCT 스펙트럼을 생성하고, 상

기 (c) 단계에서는 쇼트 윈도우(short window)를 적용하여 쇼트 FFT 스펙트럼을 생성하고, 상기 (d) 단계에서는 상기 생성된 롱 CMDCT 스펙트럼 및 쇼트 FFT 스펙트럼에 기초하여 심리 음향 모델 분석을 수행하는 것을 특징으로 하는 방법.

【청구항 5】

제1항에 있어서, 상기 (a) 단계에서 결정된 윈도우 타입이 쇼트 윈도우인 경우, 상기 (b) 단계에서는 쇼트 윈도우를 적용하여 쇼트 CMDCT 스펙트럼을 생성하고, 상기 (c) 단계에서는 롱 윈도우를 적용하여 롱 FFT 스펙트럼을 생성하고, 상기 (d) 단계에서는 상기 생성된 쇼트 CMDCT 스펙트럼 및 롱 FFT 스펙트럼에 기초하여 심리 음향 모델 분석을 수행하는 것을 특징으로 하는 방법.

【청구항 6】

제1항에 있어서, 상기 (a) 단계는 입력 오디오 신호가 천이(transient) 신호인 경우에는 윈도우 타입을 쇼트 윈도우로 결정하고, 비천이 신호인 경우에는 윈도우 타입을 롱 윈도우로 결정하는 것을 특징으로 하는 방법.

【청구항 7】

제1항에 있어서, (e) 상기 (d) 단계에서 수행된 심리 음향 모델 분석 결과에 기초하여 양자화 및 부호화를 수행하는 단계를 더 포함하는 것을 특징으로 하는 방법.

【청구항 8】

제1항에 있어서, 상기 심리 음향 모델은 MPEG(motion picture experts group)-1 레이어 3, MPEG-2 AAC(advanced audio coding), MPEG 4, WMA(windows media audio)으로 이루어진 그룹 중 어느 하나에서 사용되는 심리 음향 모델인 것을 특징으로 하는 방법.

【청구항 9】

디지털 오디오 데이터 부호화 장치에 있어서,
입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 윈도우 스위칭부와,
상기 윈도우 스위칭부에서 결정된 윈도우 타입에 따라 상기 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 CMDCT부와,
상기 윈도우 스위칭부에서 결정된 윈도우 타입을 이용하여 상기 입력 오디오 신호로부터 FFT 스펙트럼을 생성하는 FFT부와,
상기 CMDCT부에서 생성된 CMDCT 스펙트럼 및 상기 FFT부에서 생성된 FFT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 심리 음향 모델부를 포함하는 것을 특징으로 하는 장치.

【청구항 10】

제9항에 있어서, 상기 부호화 장치는 상기 입력 오디오 신호를 필터링하여 복수개의 서브 밴드로 분할하는 필터부를 더 포함하며, 상기 윈도우 스위칭부는 상기 필터부의 출력 데이터에 기초하여 윈도우 타입을 결정하는 것을 특징으로 하는 장치.

【청구항 11】

제10항에 있어서, 상기 필터부는 다상 필터 बैं크인 것을 특징으로 하는 장치.

【청구항 12】

제9항에 있어서, 상기 윈도우 스위칭부에서 결정된 윈도우 타입이 롱 윈도우인 경우, 상기 CMDCT부는 롱 윈도우를 적용하여 롱 CMDCT 스펙트럼을 생성하고, 상기 FFT부는

쇼트 윈도우를 적용하여 쇼트 FFT 스펙트럼을 생성하고, 상기 심리 음향 모델부는 상기 CMDCT부에서 생성된 롱 CMDCT 스펙트럼 및 상기 FFT부에서 생성된 쇼트 FFT 스펙트럼에 기초하여 심리 음향 모델 분석을 수행하는 것을 특징으로 하는 장치.

【청구항 13】

제9항에 있어서, 상기 윈도우 스위칭부에서 결정된 윈도우 타입이 쇼트 윈도우인 경우, 상기 CMDCT부는 쇼트 윈도우를 적용하여 쇼트 CMDCT 스펙트럼을 생성하고, 상기 FFT부는 롱 윈도우를 적용하여 롱 FFT 스펙트럼을 생성하고, 상기 심리 음향 모델부는 상기 CMDCT부에서 생성된 쇼트 CMDCT 스펙트럼 및 상기 FFT부에서 생성된 롱 FFT 스펙트럼에 기초하여 심리 음향 모델 분석을 수행하는 것을 특징으로 하는 장치.

【청구항 14】

제9항에 있어서, 상기 윈도우 스위칭부는 입력 오디오 신호가 천이 신호인 경우에는 윈도우 타입을 쇼트 윈도우로 결정하고, 비천이 신호인 경우에는 윈도우 타입을 롱 윈도우로 결정하는 것을 특징으로 하는 장치.

【청구항 15】

제9항에 있어서, 상기 CMDCT부로부터의 오디오 데이터 및 상기 심리 음향 모델부로부터의 결과값에 기초하여 양자화 및 부호화를 수행하는 양자화 및 부호화부를 더 포함하는 것을 특징으로 하는 장치.

【청구항 16】

제9항에 있어서, 상기 심리 음향 모델은 MPEG-1 레이어 3, MPEG-2 AAC, MPEG 4, WMA으로 이루어진 그룹 중 어느 하나에서 사용되는 심리 음향 모델인 것을 특징으로 하는 장치.

【청구항 17】

디지털 오디오 부호화 방법에 있어서,

(a) 입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 단계와,

(b) 상기 생성된 CMDCT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 단계를 포함하는 것을 특징으로 하는 방법.

【청구항 18】

제17항에 있어서, 상기 (a) 단계는 (a1) 입력 오디오 신호에 대해 롱 윈도우 및 쇼트 윈도우를 적용하여 CMDCT를 수행하여, 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 생성하는 단계를 더 포함하는 것을 특징으로 하는 방법.

【청구항 19】

제18항에 있어서, 상기 (b) 단계는 상기 (a1) 단계에서 생성된 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 사용하여 심리 음향 모델 분석을 수행하는 것을 특징으로 하는 방법.

【청구항 20】

제17항에 있어서, 상기 (a) 단계는 (a2) 상기 입력 오디오 신호를 필터링하여 복수 개의 서브 밴드로 분할하는 단계를 더 포함하며, 상기 CMDCT 스펙트럼을 생성하는 단계는 상기 서브 밴드로 분할된 데이터에 대해 수행되는 것을 특징으로 하는 방법.

【청구항 21】

제17항에 있어서, 상기 부호화 방법은 (a3) 상기 입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 단계를 더 포함하는 것을 특징으로 하는 방법.

【청구항 22】

제21항에 있어서, 상기 (a3) 단계는 입력 오디오 신호가 천이 신호인 경우에는 윈도우 타입을 쇼트 윈도우로 결정하고, 비천이 신호인 경우에는 윈도우 타입을 롱 윈도우로 결정하는 것을 특징으로 하는 방법.

【청구항 23】

제20항에 있어서, 상기 (a2) 단계는 다상 필터 뱅크에 의해 수행되는 것을 특징으로 하는 방법.

【청구항 24】

제22항에 있어서, 상기 (a3) 단계에서 결정된 윈도우 타입이 롱 윈도우인 경우에는, 상기 (b) 단계에서 수행된 심리 음향 모델 분석 결과에 기초하여 롱 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하고, 상기 (a2) 단계에서 결정된 윈도우 타입이 쇼트 윈도우인 경우에는, 상기 (b) 단계에서 수행된 심리 음향 모델 분석 결과에 기초

하여 쇼트 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하는 단계를 포함하는 것을 특징으로 하는 방법.

【청구항 25】

제17항에 있어서, 상기 심리 음향 모델은 MPEG-1 레이어 3, MPEG-2 AAC, MPEG 4, WMA로 이루어진 그룹 중 어느 하나에서 사용되는 심리 음향 모델인 것을 특징으로 하는 방법.

【청구항 26】

디지털 오디오 부호화 장치에 있어서,

입력 오디오 신호로부터 CMDCT 스펙트럼을 생성하는 CMDCT부와,

상기 CMDCT부에서 생성된 CMDCT 스펙트럼을 이용하여 심리 음향 모델 분석을 수행하는 심리 음향 모델부를 포함하는 것을 특징으로 하는 장치.

【청구항 27】

제26항에 있어서, 상기 CMDCT부는 상기 입력 오디오 신호에 대해 롱 윈도우 및 쇼트 윈도우를 적용하여 CMDCT를 수행하여, 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 생성하는 것을 특징으로 하는 장치.

【청구항 28】

제27항에 있어서, 상기 심리 음향 모델부는 상기 CMDCT부에서 생성된 롱 CMDCT 스펙트럼 및 쇼트 CMDCT 스펙트럼을 사용하여 심리 음향 모델 분석을 수행하는 것을 특징으로 하는 장치.

【청구항 29】

제26항에 있어서, 상기 입력 오디오 신호를 필터링하여 복수개의 서브 밴드로 분할하는 필터부를 더 포함하며, 상기 CMDCT부는 상기 서브 밴드로 분할된 데이터에 대해 CMDCT를 수행하는 것을 특징으로 하는 장치.

【청구항 30】

제26항에 있어서, 상기 입력 오디오 신호의 특성에 따라 윈도우 타입을 결정하는 윈도우 타입 결정부를 더 포함하는 것을 특징으로 하는 장치.

【청구항 31】

제30항에 있어서, 상기 윈도우 타입 결정부는 입력 오디오 신호가 천이 신호인 경우에는 윈도우 타입을 쇼트 윈도우로 결정하고, 비천이 신호인 경우에는 윈도우 타입을 롱 윈도우로 결정하는 것을 특징으로 하는 장치.

【청구항 32】

제29항에 있어서, 상기 필터부는 다상 필터 뱅크인 것을 특징으로 하는 장치.

【청구항 33】

제31항에 있어서, 상기 부호화 장치는 양자화 및 부호화부를 더 포함하며, 상기 양자화 및 부호화부는 상기 윈도우 타입 결정부에서 결정된 윈도우 타입이 롱 윈도우인 경우에는, 상기 심리 음향 모델부에서 수행된 심리 음향 모델 분석 결과에 기초하여 롱 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하고, 상기 윈도우 타입 결정부에서 결정된 윈도우 타입이 쇼트 윈도우인 경우에는, 상기 심리 음향 모델부에서 수행된 심리 음

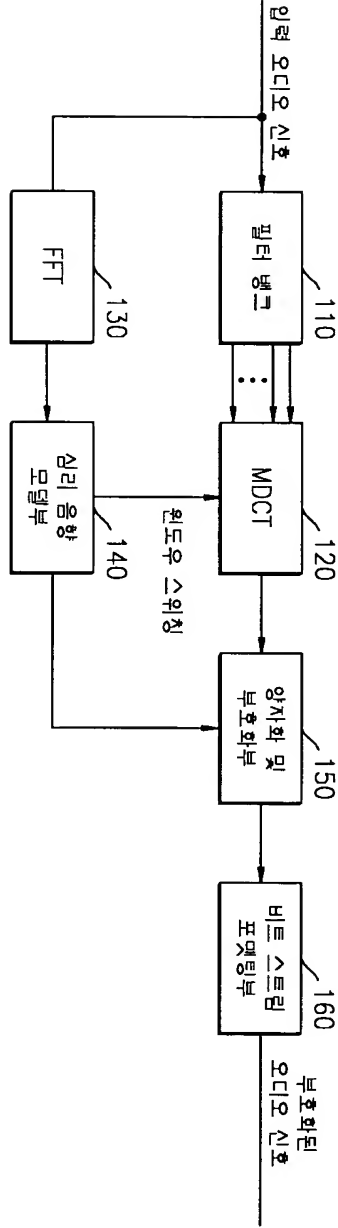
향 모델 분석 결과에 기초하여 쇼트 MDCT 스펙트럼에 대해 양자화 및 부호화를 수행하는 것을 특징으로 하는 장치.

【청구항 34】

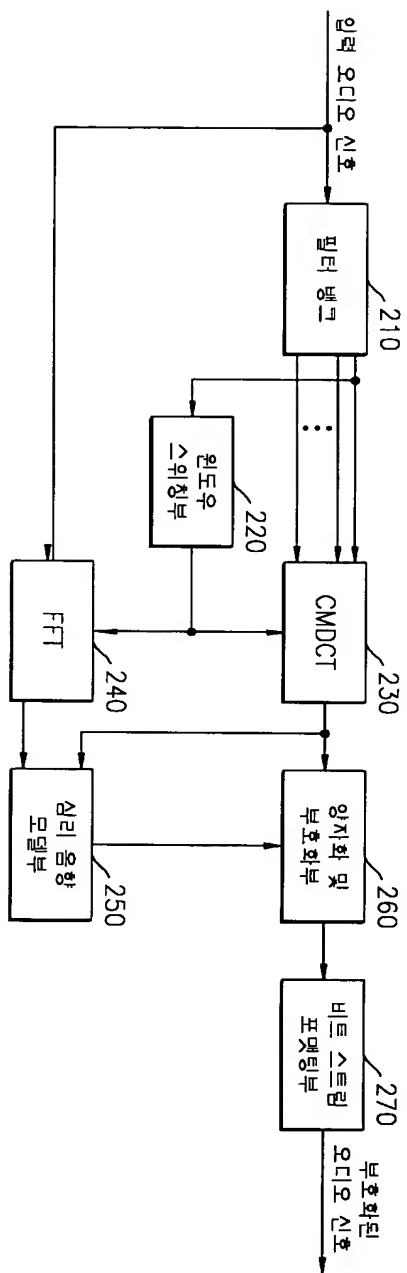
제26항에 있어서, 상기 심리 음향 모델은 MPEG-1 레이어 3, MPEG-2 AAC, MPEG 4, WMA로 이루어진 그룹 중 어느 하나에서 사용되는 심리 음향 모델인 것을 특징으로 하는 장치.

【도면】

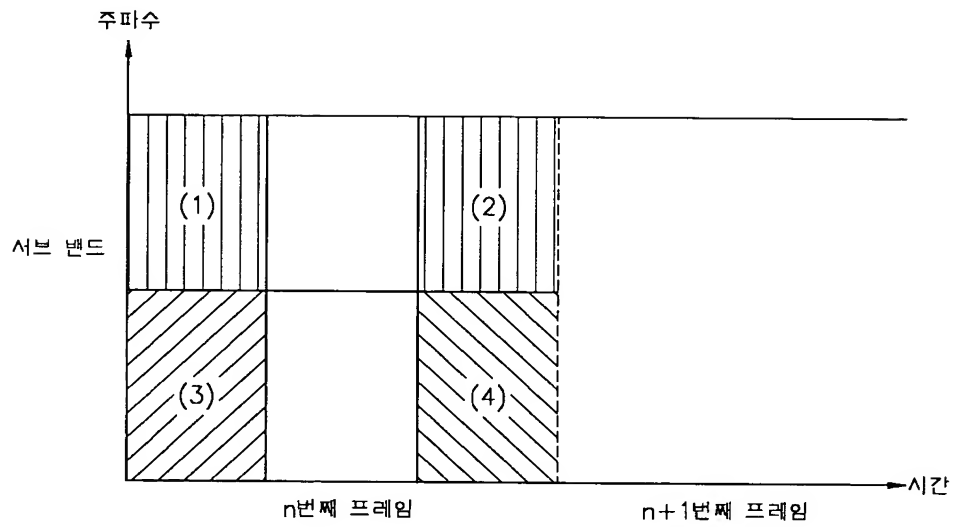
【도 1】



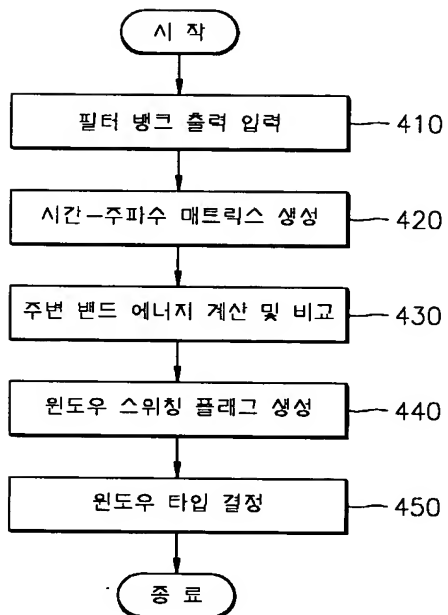
【도 2】



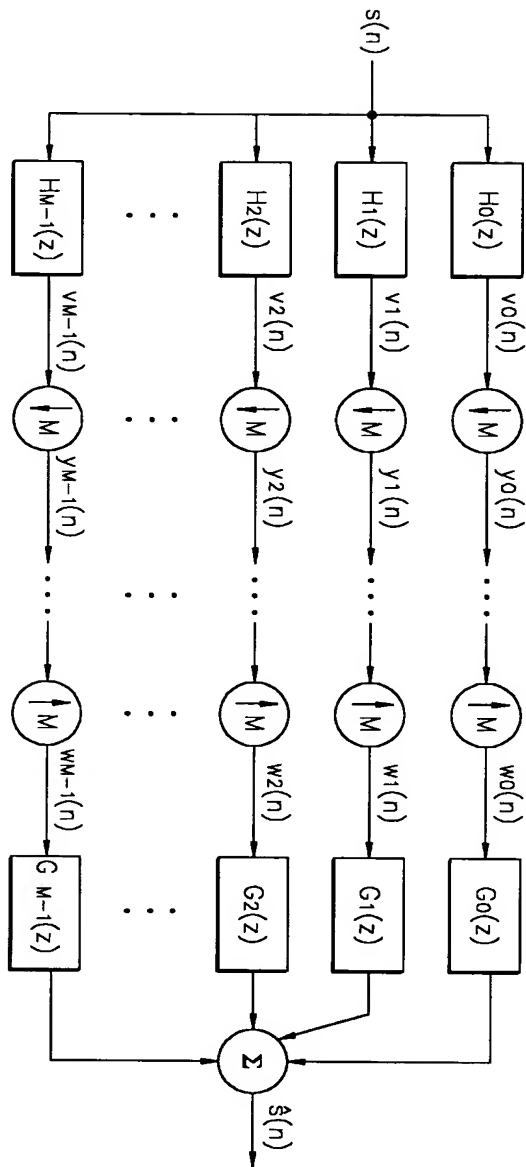
【도 3】



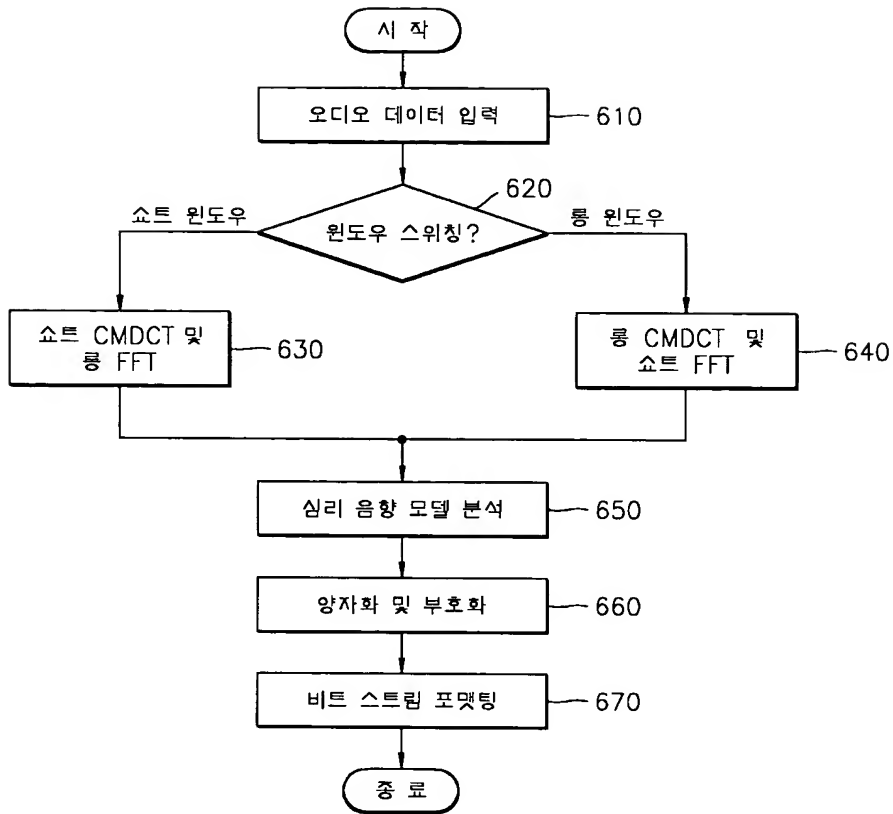
【도 4】



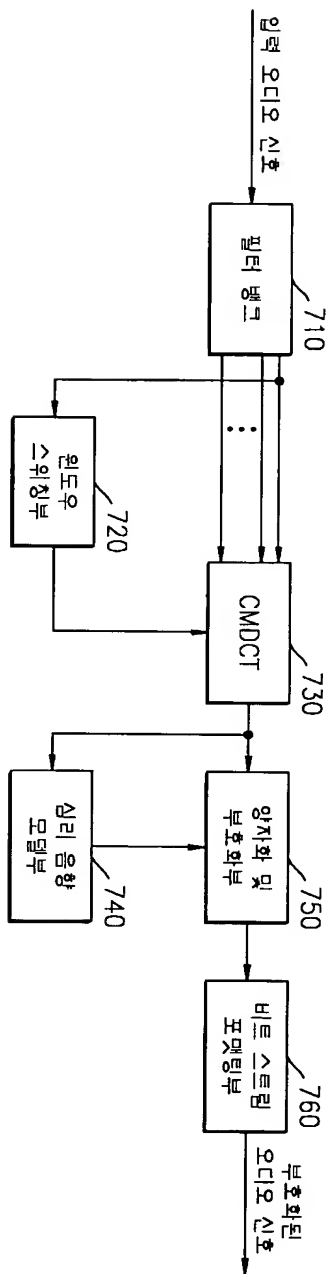
【도 5】



【도 6】



【도 7】



【도 8】

